

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-055185

(43)Date of publication of application : 27.02.1996

(51)Int.Cl.

G06K 9/20
G06K 9/62

(21)Application number : 06-191408

(71)Applicant : MATSUSHITA ELECTRIC IND CO LTD

(22)Date of filing : 15.08.1994

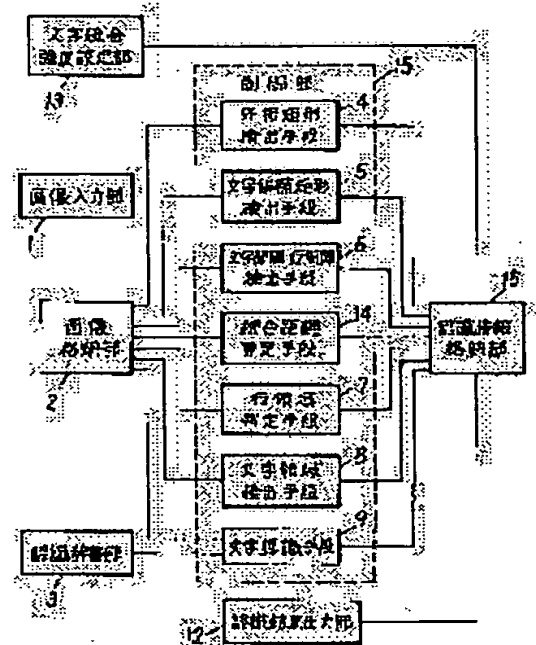
(72)Inventor : OGAWA JUN

(54) CHARACTER RECOGNITION DEVICE

(57)Abstract:

PURPOSE: To provide a character recognition device which accurately and speedily decides a line area and is high in the reliability of character recognition.

CONSTITUTION: The character recognition device is equipped with a character integration strength setting part 13 which sets a reference value for deciding a line area, an integration distance calculating means 14 which calculates an integration distance by multiplying the most frequent value of character intervals detected by a character interval and line interval detecting means 6 by the reference value set by the character integration strength setting part 13, and a recognition information storage part 16 which stores the reference value set by the character integration strength setting part 13 and the integration distance calculated by the integration distance calculating means 14.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

【特許請求の範囲】・

特開平8-551855

(43) 公開日 平成8年(1996)2月27日

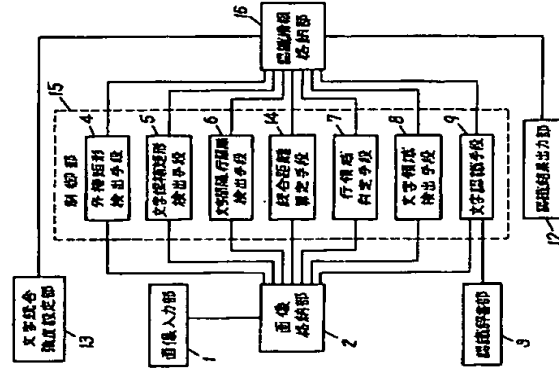
(51) Int. Cl. ⁸	識別記号	序内整理番号	FI	技術表示場所
G 0 6 K	9/20 3 4 0 K			
	9/62 6 1 0 Z	9061-5 H		
登録請求 未請求 請求項の数 1			OL	(全 8 頁)
(21) 出願番号	特願平6-191408		(71) 出願人	000005821 松下電器産業株式会社
(22) 出願日	平成6年(1994)8月15日		(72) 発明者	小川 潤 大阪府門真市大字門真1006番地
			(74) 代理人	小堀 治 明 (外2名) 産業株式会社内 弁理士

(54)【發明の名称】文字認識装置

(57) 【题解】

【目的】 行領域の判定の正確で迅速な、かつ文字認識の信頼性の高い文字認識装置を提供することを目的とする。

【備考】 文字距離算定であって、行間値を判定する基準となる文字結合強度設定部13と、文字間隔、行間隔抽出手段6により抽出された文字間隔の最頻値に文字結合強度設定部13で設定された基準値を掛け合わせた該強度を算定する統合距離算定手段14と、文字結合強度設定部13で設定された該強度、及び、統合距離算定手段14により算定された該強度を格納している距離情報記憶部16と、を備えた構成を有する。



【統計請求の範囲】・

【請求項1】 文字データを含む画像データを入力する画像入力部と、前記画像入力部に入力された前記画像データの

[illegible]

【発明の詳細な説明】

【0001】
【産業上の利用分野】本発明は印刷文書のデータベータス
化や文書の再利用のために、スキャナ等の光学的手段を
用いて文書画像を取り込み、取り込んだ画像データから
文字、図形、装飾に属性毎に領域を抽出し、各属性に付
した認識処理を行う文字認識装置に関するものである。

[0002]

【従来の技術】近年、文書等の文字、図形、表等の情報を含む画像データを文字認識装置で処理することが多くおこなわれている。この文字認識装置を用いて、画像データで表されている文字情報を「I S Oコード等のコード情報（文字認識装置では認識結果と呼んでいる）に変換して処理されている。

【0003】以下に従来の文字認識装置について説明す 50

ク図である。図7において、1は文字データを含む画像データを入力する画像入力部、2は画像入力部1に入力された画像データを格納する画像格納部、3は文字の図形特徴を格納する図形特徴部、4は画像格納部1に格納された画像データから照像素領域に外接する外接矩形を抽出する外接矩形抽出手段、5は外接矩形抽出手段4により抽出された外接矩形から文字候補矩形を抽出する文字候補矩形抽出手段、6は文字候補矩形抽出手段5による

10 り抽出された文字設備短形間の文字間隔及び行間隔を算出する文字間隔、行間隔抽出手段、7は文字間隔、行間隔抽出手段6で抽出した文字間隔が予め定められている統合回数より短い場合は同一の行間隔であると判定する行間隔判定手段、8は行間隔判定手段7により判定された行間隔から文字設備を抽出する文字設備抽出手段、9は文字設備抽出手段8により抽出された文字領域と図解手帳部3に格納された図形特徴とを照合することにより図解結果を得る文字認識手段、10は外接短形抽出手段4と、文字設備短形抽出手段5と、文字間隔、行間隔抽出手段6と、行間隔判定手段7と、文字設備抽出手段8と、文字認識手段9とを有する制御部、11は外接短形抽出手段4により抽出された外接短形と、文字設備短形抽出手段5により抽出された文字設備短形と、文字間隔抽出手段6により抽出された文字間隔及び行間隔と、行間隔判定手段7により判定された行間隔と、文字設備抽出手段8により抽出された文字領域と、文字認識手段9において照合された文字認識結果とを格納する認識情報格納部、12は文字認識手段9において照合された認識結果を出力する認識結果出力部である。

[illegible]

いて、以下その動作を説明する。まず、文字データを各
 画面データとスキャナ等の画面入力部11に入力し画像
 格納部2に格納する。次に、外縁矩形抽出手段4により
 画面格納部2に格納された2値データの画面裏が逆転し
 ていたかたまりに外接する矩形（以下外縁矩形と呼ぶ）

の要領を抽出し、配属情報格納部11に格納する。次に、文字候補矩形抽出手段5において外接矩形抽出手段4により抽出した外接矩形が比較的大きな矩形を図形候補矩形、それ以外を文字候補矩形と抽出し、図形候補矩形と文字候補矩形の要領を配属情報格納部11に格納する。次に、文字間隔、行間隔を抽出手段6により文字候補矩形の水平方向と垂直方向の間隔を測定し、文字間隔と行間隔を抽出し、文字間隔と行間隔を配属情報格納部11に格納する。

1に格納する。次に、行領域判定手段7により文字間
隔、行間隔検出手段8により検出された文字間隔の最
大値に固定値である被合距離をかけた値を基調として、こ
の値より文字間隔が小さなものは同一行領域と判定し、
行領域を認識情報格納部11に格納する。次に、文字検
出手段9により行領域判定手段7により判定された行
領域から文字間隔を検出し、認識情報格納部11に格
納する。

8

を示す機能ブロック図

【図2】本発明の一実施例における文字認識装置の構成

を示す装置ブロック図

【図3】本発明の一実施例における文字認識装置の動作

を示すフローチャート

【図4】本発明の一実施例における文字統合強度設定の

CRT画面を示す図

【図5】本発明の一実施例における画像データの例を示

す図

【図6】本発明の一実施例における図5で示した画像デ

ータの外接矩形を示す図

【図7】従来の文字認識装置の構成を示す構成ブロック

図

【図8】従来の文字間隔が大きい場合の行領域の判定を

示す図

【図9】従来の文字間隔が小さい場合の行領域の判定を

示す図

【符号の説明】

1 画像入力部

2 画像格納部

3 認識評価部

4 外接矩形検出手段

(5)

5 文字候補矩形検出手段

6 文字間隔、行間隔検出手段

7 行領域判定手段

8 文字領域検出手段

9 文字認識手段

10 制御部

11 認識情報格納部

12 認識結果出力部

13 文字統合強度設定部

14 統合距離算定手段

15 制御部

16 認識情報格納部

17 スキャナ

18 CRT

19 キーボード

20 マウス

21 中央処理装置

22 画像格納部

23 認識評価部

24 認識情報格納部

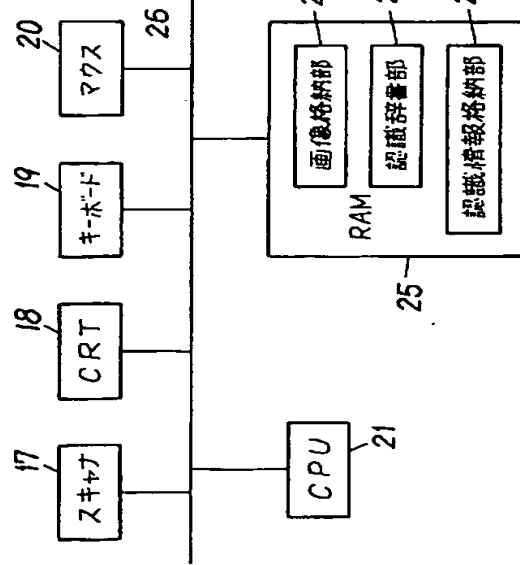
25 ランダム・アクセス・メモリ

26 システム・バス

【図2】

【図5】

今日は、天気が
よい

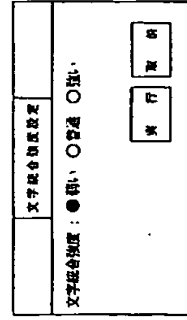


【図6】

今日は、天気が
よい

【図4】

【図8】



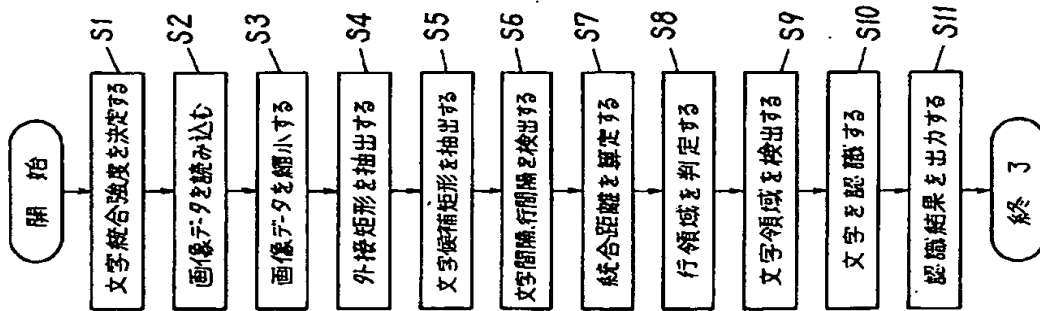
今日は、天気がよいので
遊びに行きます。

【図9】

今日は、天気がよいので
遊びに行きます。

今日は、天気がよいので
遊びに行きます。

【図3】



【図7】

